Acceptance, Truth, and Probability: The Case of Conditionals[1]

Ginger Schultheis

August 2023

1 Introduction

In her classic paper *On Conditionals*, Dorothy Edgington reminds us of the importance of testing the predictions that a semantic theory of conditionals makes about our *less-than-certain judgments*—our non-extreme credences—in conditionals. If we find that our credences in conditionals are radically different from those we would expect to have if we were forming credences in accordance with a certain theory, then that is evidence against the theory. If, moreover, there is an alternative theory that makes better predictions about our less-than-certain judgments, then that is reason to prefer the alternative.

This is, more or less, the state of play as I see it in current theorizing about conditionals. There are two leading theories—the *strict theory* and the *variably strict theory*. Our ordinary credences in conditionals are far out of line with what the strict theory recommends. And there's a variably strict theory that does better: Stalnaker's variably strict theory.[2] In light of this, I recommend that we reject the strict theory in favor of a Stalnakerian variably strict theory.

But certain inference patterns for both indicatives and counterfactuals that are valid on the strict theory and invalid on the variably strict theory—such as Transitivity, Contraposition, and Antecedent Strengthening—do share certain good-making features with classically valid arguments. Transitivity and Contraposition certainly seem like excellent principles. Although Antecedent Strengthening may seem less obvious at first, strict theorists have convincingly argued that the full range of data surrounding this principle is well explained by contextualist or dynamic strict theories.

My task in this paper will be to develop a Stalnakerian variably strict theory on which the characteristically strict inference patterns are *pseudo-valid*: roughly, if the premises can be felicitously asserted in a given context, then if the premises are known in that context, the conclusion is also true in the context.[3] The idea that these inferences for *indicative conditionals* have some good

[1]Thanks to Andrew Bacon, David Boylan, Cian Dorr, Melissa Fusco, Matthew Mandelkern, Arc Kocurek, Justin Khoo, Paolo Santorio, and Snow Zhang for helpful conversations about this material. Special thanks to Milo Phillips-Brown for feedback on multiple drafts.

[2]See Stalnaker (1968) and Stalnaker (1975).

[3]Pseudo-validity is closely related to the notion of *quasi-validity* in Kolodny and Macfarlane (2010) and Dorr & Hawthorne (ms). See footnote 20.

status like pseudo-validity is already nascent in Stalnaker's theory of indicatives. I develop this thought in more detail in §5. But the parallel question of whether these inferences *for counterfactuals* are pseudo-valid—or have some other good status that falls short of classical validity—has gone unexplored. In §6-§7, I develop a new Stalnakerian theory of counterfactuals, and I show that the theory predicts that our three inference patterns are pseudo-valid.

Before I present these theories, I start in §2 by presenting the strict theory and the variably strict theory. In §3, I present the credence-theoretic argument against the strict theory and explain why Stalnaker's theory does better. In §4, I present the case for the characteristically strict inference patterns: Transitivity, Contraposition, and Antecedent Strengthening.

2 Strict and Variably Strict Conditionals

According to the strict theory, natural language conditionals are necessitated material conditionals.[4] The flavor of necessity depends on the kind of conditional. For indicatives, it is epistemic necessity. An indicative conditional $\ulcorner A > B \urcorner$ is true if and only if $\ulcorner A \supset B \urcorner$ is true throughout the epistemically possible worlds. For counterfactuals, the necessity is historical or metaphysical necessity. A counterfactual $\ulcorner A \mathrel{\Box\!\!\rightarrow} B \urcorner$ is true just in case $\ulcorner A \supset B \urcorner$ is true throughout the historically or metaphysically possible worlds. Let W be a non-empty set of worlds. Let R_c be a contextually-supplied, reflexive accessibility relation over W. Letting '$\rightarrow$' stand for both indicatives and counterfactuals, we have:

Strict Theory
$[\![A \rightarrow B]\!]^{c,w} = 1$ if and only if $\mathsf{R}_c(w) \cap [\![A]\!]^c \subseteq [\![B]\!]^c$

The strict theory retains many classical properties of the material conditional. Consider:

Transitivity
$A \rightarrow B, B \rightarrow C \vDash A \rightarrow C$

Contraposition
$A \rightarrow B \vDash \neg B \rightarrow \neg A$

Antecedent Strengthening
$A \rightarrow B \vDash (A \wedge C) \rightarrow B$

[4] See Warmbrod (1981), Veltman (1985), von Fintel (2001), and Gillies (2007) for defenses of strict theories.

All three inferences are valid on the Strict Theory.

Before moving on, a qualification. It is standard to say that conditionals carry a *compatibility presupposition*. A conditional presupposes that there are accessible antecedent-worlds. Often this presupposition is formalized in a trivalent framework—sentences whose presuppositions are not satisfied are neither true nor false. On these theories, Antecedent Strengthening and Contraposition are merely *Strawson-valid*: if the premises are true and the conclusion is either true or false, the conclusion is true.[5] Others prefer a *multidimensional* treatment of presupposition: sentences are always true or false, and presupposition is an independent dimension of meaning.[6] On this theory, the inferences are classically valid. For ease of exposition, I'll assume a multidimensional treatment.

To state the variably strict theory, we introduce a *selection function* f_c that takes a world and an antecedent proposition to a set of worlds such that:

$$f_c(w, [\![\mathrm{A}]\!]^c) \subseteq \mathsf{R}_c(w) \cap [\![\mathrm{A}]\!]^c \tag{1}$$

Then:

Variably Strict Theory
$[\![\mathrm{A} \rightarrow \mathrm{B}]\!]^{c,w} = 1$ if and only if $f_c([\![\mathrm{A}]\!]^c, w) \subseteq [\![\mathrm{B}]\!]^c$

The variably strict theory invalidates Transitivity, Contraposition, and Antecedent Strengthening.[7]

3 Less-than-Certain Judgments

I turn now to our less-than-certain judgments in conditionals. Our ordinary, non-extreme credences in conditionals differ drastically from those predicted by the strict theory. In contrast, Stalnaker's variably strict theory fares much better. In this section, I bring this matter into view, first showing the strict theory's struggles with our less-than-certain judgments (3.1) and then Stalnaker theory's success with them (3.2).

3.1 The Strict Theory

Start with an observation due to Edgington (1995). We often find ourselves uncertain of a conditional even though we know that not all accessible antecedent-worlds are consequent-worlds. Take an example. There are twenty bikes at my

[5] See von Fintel (2001).

[6] Herzberger (1973) and Kartunnen and Peters (1979).

[7] For variably strict theories, see Stalnaker (1968) and Lewis (1973).

gym. Usually, but not always, if I get to the gym by three, at least one is free. I say:

(1) I'm confident that if I get to the gym by three, one of the bikes will be free.

Here I express high confidence in the proposition expressed by:

(2) If I get to the gym by three, one of the bikes will be free.

Now, the Strict Theory says that (2) is true if and only if the material conditional

(3) Either I do not make it to the gym by three or one of the bikes will be free.

is epistemically necessary. But my credences suggest otherwise. I assign high credence to (2). But I know that (3) is not epistemically necessary—I know I that *might* make it to the gym by three and find no free bikes.

Here's a case involving counterfactuals. John and I are talking about the NBA Finals. Last year the Warriors faced the Celtics, and the Warriors won. John asks:

(4) What's the chance the Warriors would have won if they'd faced the Bucks?

I reply:

(5) About 20-30%, I'd say.

Here I express low, non-zero credence in the counterfactual:

(6) If the Warriors had faced the Bucks, they would have won.

The strict theory says that (6) is true if and only if (7) is historically or metaphysically necessary.

(7) Either the Warriors did not face the Bucks or they won.

My credences again suggest otherwise. I assign high credence to the proposition expressed by (6). But I know that (7) is not historically or metaphysically necessary—I know that the Warriors *could have* played the Bucks and lost.

For a second argument against the strict theory on the basis of our less-than-certain judgments, consider:

Conditional Excluded Middle
$\vDash A \rightarrow B \vee A \rightarrow \neg B$

As many authors observe, our credences tend to conform to Conditional Excluded Middle.[8] Notice that insofar as I have low confidence in (6), repeated below,

(6) If the Warriors had faced the Bucks, they would have won.

[8] See van Fraassen (1976), Santorio (2017), Mandelkern (2019), and Dorr & Hawthorne (ms).

I should have *high* confidence in (8):

(8) If the Warriors had faced the Bucks, they would have lost.

Since I'm around 20-30% confident in (6), I should be 70-80% confident in (8). Furthermore, I assign zero credence to the conjunction:

(9) If the Warriors had faced the Bucks in the Finals, they would have won, and if they had faced the Bucks in the Finals, they would have lost.

If my credences obey the axioms of probability, then it follows that I am certain of the following instance of Conditional Excluded Middle.

(10) Either, if the Warriors had faced the Bucks in the Finals, they would have won, or if they had faced the Bucks in the Finals, they would have lost.

Consider a second example from Mandelkern (2019). Mark is holding a coin.

(11) [Jack:] If Mark flips the coin, it will land heads.

(12) [Sue:] If Mark flips the coin, it will land tails.

If you think the coin is fair, you should be 50% confident in what Jack says and 50% confident in what Sue says. If you learn the coin is weighted towards heads at a ratio of 3:1, you should be 75% confident in what Jack says and 25% confident in what Sue says. Likewise, if you learn that the ratio is 4:1, then you should increase your credence in (11) to 80% and decrease your credence in (12) to 20%.

This pattern is pervasive. Our credences in $\ulcorner A \to B \urcorner$ and $\ulcorner A \to \neg B \urcorner$ tend to sum to one. Since we ordinarily assign zero credence to their conjunction, it follows that we are conforming to the principle of Conditional Excluded Middle.

But Conditional Excluded Middle is not valid on the strict theory. Whenever there are accessible worlds where $\ulcorner$A and B$\urcorner$ is true, and accessible worlds where $\ulcorner$A and $\neg$B$\urcorner$ is true, Conditional Excluded Middle fails. On the strict theory, it is a mystery why competent form credences in accordance with Conditional Excluded Middle.

3.2 Stalnaker's Theory

We have seen that our ordinary credences in conditionals are at odds with the strict theory's predictions. This argument wouldn't show much if it turned out that no theory does any better. But there is a theory that does better—a theory that predicts and explains our patterns of credence formation in conditionals, both indicative and counterfactual. That theory is Stalnaker's variably strict theory.[9]

Stalnaker says that a conditional $\ulcorner A \rightarrow B \urcorner$ is true, at a world w, if and only if the closest A-world, to w, is a B-world. (I'll give a more precise statement in §5.) Stalnaker's Theory validates Conditional Excluded Middle. Either the closest A-world is a B-world, or the closest A-world is a ¬B-world. In the first case, $\ulcorner A \rightarrow B \urcorner$ is true; in the second, $\ulcorner A \rightarrow \neg B \urcorner$ is true. When it comes to our less-than-certain judgments, this is already significant progress. If Conditional Excluded Middle is valid, we should expect competent speakers to form credences in accordance with this principle, and that is exactly what we find.

We can go further. We can give a much more general credence-theoretic argument in favor of Stalnaker's theory. For we can ask: what are our credences in conditionals, quite generally? A compelling and widely-accepted answer to this question in the case of indicative conditionals is Stalnaker's Thesis, stated below.

> **Stalnaker's Thesis**
> The probability that you assign to an indicative conditional is equal to the probability that you assign to its consequent conditional on its antecedent.

Take an example. I'm holding a fair, six-sided die in my hand. Consider:

(13) If I roll the die, it will land on one.

Intuitively, the probability of (13) is 1/6. That is also the conditional probability that the die lands on one, given that I roll the die. It is easy to multiply examples like this. In general, we calculate the probability of an indicative conditional $\ulcorner A > B \urcorner$ by calculating the probability of B conditional on A. That is just what we would expect if Stalnaker's Thesis were true.

Building on the work of van Fraassen (1976), various authors, such as Kaufman (2009), and Bacon (2015), have proven *tenability results* showing that certain versions of Stalnaker's Thesis hold on certain versions of Stalnaker's Theory. Though the details are complex, we can give an informal explanation of how these results work.[10]

Stalnaker imposes four constraints on the selection function for indicative conditionals. First: the selected A-world is always an A-world. Second: if w is an A-world, then the selected A-world, at w, is w itself. Third: if there are epistemically possible A-worlds, then the selected A-world is epistemically possible. Fourth: if there are no epistemically possible worlds, then there is no selected A-world: $f_c(w, [\![A]\!]^c) = \emptyset$.

[9] Stalnaker (1968) and Stalnaker (1975).
[10] For this informal explanation, I am indebted to Schulz (2017).

Assume these are the only constraints.[11] Then, if we're at an A-world, the selected A-world is the actual world. Otherwise, we can think of the selection function as randomly selecting a world from the epistemically possible A-worlds. With this metaphor in mind, we can begin to see how Stalnaker's semantics can capture our less-than-certain judgments. For one, we can be confident that a randomly selected A-world is a B-world even if we know that not all epistemically possible A-worlds are B-worlds. And our credences in conditionals can be sensitive to the distribution of consequent-worlds among the epistemically possible antecedent-worlds. If few epistemically possible A-worlds are B-worlds—the probability of B given A is low—then the randomly selected A-world will probably not be a B-world, and so the probability of the conditional will be low. If most epistemically possible A-worlds are B-worlds—the probability of B given A is high—then the selected A-world will probably be B-world, and so the probability of the conditional will be high.[12]

I have been focusing on indicatives, but there is also reason to be optimistic that a Stalnakerian theory of counterfactuals can make reasonable predictions about our credences in counterfactuals.[13]

Let's take stock. Stalnaker's theory has a clear advantage over the strict theory when it to our less-than-certain judgments in conditionals. I therefore recommend rejecting the strict theory, and Transitivity, Contraposition, and Antecedent Strengthening along with it. Nevertheless, as I said at the outset, these inferences do share certain good-making features with classically valid inferences. In the next section, I make that case.

4 The Inferences

We'll start with Transitivity and Contraposition, and then we'll turn to Antecedent Strengthening.

Transitivity and Contraposition seem to be fundamental principles in our ordinary reasoning with conditionals. Milo and I are talking about seeding for the NBA playoffs. I have reason to believe:

(14) If the Lakers win tonight, they will secure the fifth seed.

Milo adds:

[11]Note that Stalnaker himself does not assume that these are the only constraints on the selection function. For example, he suggests that the selection function may have a preference for *more similar* worlds. But as Bacon (2015) convincingly explains, we must abandon this idea if we want to capture Stalnaker's Thesis.

[12]The idea of randomness is made explicit in Schulz (2014) and Bacon (2015).

[13]For recent work on this topic, see Khoo (2022), Schulz (2017), and Schultheis (forthcoming).

(15) If the Lakers secure the fifth seed, they will play the Warriors.

I will and I should put these two together, concluding:

(16) If the Lakers win tonight, they will play the Warriors.

The inference is no less compelling with counterfactuals. The playoffs have finished. Milo and I are talking about how things might have gone if the Lakers had won that night. Milo says:

(17) If the Lakers had won, they would have secured the fifth seed.

(18) If they had secured the fifth seed, they would have played the Warriors.

I will and I should conclude:

(19) If the Lakers had won, they would have played the Warriors.

Contraposition also appears to be an excellent form of inference. You tell me:

(20) If the ring is made of gold, it won't turn green.

I conclude:

(21) If it turns green, it isn't made of gold.

Or suppose I know:

(22) If Durant played the whole first half, he didn't play the whole second.

Then I have every reason to conclude:

(23) If Durant did play the whole second half, he didn't play the whole first.

Contraposition is much less commonly used with counterfactuals. But I don't think this alone gives us reason to doubt its validity. Counterfactuals tend to implicate that their antecedents and consequents are false, and so it's hard to find contexts in which $\ulcorner A \mathrel{\Box\!\!\rightarrow} B \urcorner$ and $\ulcorner \neg B \mathrel{\Box\!\!\rightarrow} \neg A \urcorner$ are both assertable. Suppose I say:

(24) If it had been made of gold, it would have turned green.

(24) strongly suggests that the ring is not made of gold and that it did not turn green. But then it would be strange to continue with

(25) If it hadn't turned green, it wouldn't have been made of gold.

since (25) strongly suggests that the ring *did* turn green.

There are, however, cases where counterfactuals do not carry this implicature, and in many of these cases Contraposition does seem like a good inference. Consider *future-less-vivid conditionals*. The ring will either turn green or red, depending on whether it is made of gold. I say:

(26) If it were made of gold, it would turn green by tomorrow.

You infer:

(27) If it were to turn red, it couldn't be made of gold.

This seems like a good inference. And note that (28) sounds incoherent:

(28) # If it were made of gold, it would turn green by tomorrow. But if it were to turn red, it could (still) be made of gold.

So far I have argued in favor of Transitivity and Contraposition. What about Antecedent Strengthening? Variably strict theorists say that it is subject to counterexample. Consider the following *Sobel sequence*.[14]

(29) If Alice comes to the party, it will be fun.

(30) But if Alice and Matt come to the party, it won't be fun.

This sequence is unremarkable. But if Antecedent Strengthening were valid, (29) and (30) would be inconsistent.

In response, strict theorists observe that Sobel sequences sound much worse when their order is reversed.[15]

(30) If Alice and Matt come to the party, it won't be fun.

(29) ? But if Alice comes to the party, it will be fun.

We are much less inclined to accept (29) after (30).

Where does this leave us with respect to the status of Antecedent Strengthening? The strict theorist hypothesizes that there is an illicit context shift in the forward Sobel sequence, and so it is not a counterexample to Antecedent Strengthening. When I encounter (29) I'm ignoring the possibility that Matt comes to the party; (30) forces me to consider this possibility. Reverse Sobel sequences, on the other hand, are usually infelicitous because the context tends not to change when the sentences are uttered in reverse order. If we encounter (30) first, we start out entertaining the possibility that Matt comes with Alice, and this possibility remains salient when we confront (29). This explanation is one that everyone should find plausible—it is well known that it is easier to widen the set of relevant possibilities than it is to narrow it.[16]

Variably strict theorists also say that Transitivity and Contraposition are subject to counterexample. For example, Stalnaker (1986) offers the following apparent counterexample to Transitivity:

---

[14] See Lewis (1973).

[15] See Warmbrod (1981), Wright (1983), Lowe (1990), von Fintel (2003), and Gillies (2007).

[16] Lewis (1979). My explanation of Sobel seqences follows von Fintel (2001) and Gillies (2007).

(31) If Hoover had been a communist, he would have been a traitor.

(32) If Hoover had been Russian, he would have been a communist.

(31) and (32) seem true. If Transitivity were valid, they would license:

(33) ? If Hoover had been Russian, he would have been traitor.

But (33) is dubious. For an apparent counterexample to Contraposition, consider a case from Adams (1988):

(34) If it rained, it didn't pour.

If Contraposition were valid, (34) would entail:

(35) ? If it poured, it didn't rain.

But (35) is false.

Again strict theorists point out that the examples exhibit signs of illicit context shifting. Consider what happens when we reverse the order of the premises in Stalnaker's example:

(32) If Hoover had been born in Russia, he would have been a communist.

(31) ? If Hoover had been a communist, he would have been a traitor.

We are much less inclined accept both premises when (31) comes after (32). This is exactly what we would expect if Stalnaker's example involved an illicit context shift.[17] Similar things can be said about the counterexample to Contraposition, but for reasons of space I won't go through the details.

Let me summarize. I agree with the strict theorist that, in general, Transitivity and Contraposition are excellent forms of reasoning; it would be a mistake to simply declare them invalid, and leave it at that. And I am not persuaded by the alleged counterexamples to any of the inferences. We have good reason to believe that they involve context-shifting.

5 Stalnakerian Indicative Conditionals

As we have just seen, a good theory of conditionals ought to explain the attractive features of Transitivity, Contraposition, and Antecedent Strengthening. I will show how to account for these inference patterns in terms of *pseudo-validity*. I start with indicatives in the §5 and turn to counterfactuals in §6.

In §5.1 I present a version of Stalnaker's theory of indicatives.[18] In §5.2, I

[17]See Lowe (1995).

[18]This version of Stalnaker's theory is not original to me. It is very similar to Bacon (2015)'s theory of indicatives, as well to a more recent theory due to Dorr & Hawthorne (ms).

show that the theory predicts that our three inferences are pseudo-valid.

5.1 The Semantics

Begin with a contextually-supplied epistemic accessibility relation E. $\mathsf{E}(w)$ is the set of worlds consistent with what's known, in w, by the conversational participants at the time of utterance. E is reflexive. Stalnaker assumes—as will I—that indicatives carry a compatibility presupposition.

> **Compatibility Presupposition**
> An indicative conditional ⌜A > B⌝ presupposes that $\mathsf{E}(w) \cap [\![A]\!]^{\mathsf{E}} \neq \emptyset$

In general, it's not appropriate to assert an indicative if you know its antecedent is false. If I assert (36), I can continue with (37), but not with (38).

(36) John is not at the party.

(37) ✓ If John had come to the party, he would have given a speech.

(38) # If John is at the party, he will give a speech.

I assume indicatives are interpreted relative to the same accessibility relation as epistemic modals. And I assume a standard relational semantics for modals:

> **Epistemic Modals**
> $[\![\text{Must A}]\!]^{\mathsf{E},w} = 1$ if and only if $\mathsf{E}(w) \subseteq [\![A]\!]^{\mathsf{E}}$

A *Stalnakerian selection function* f_{E} is a contextually-supplied function that takes a world and a proposition to a set containing at most one world. Then:

> **Stalnaker Semantics for Indicatives**
> $[\![A > B]\!]^{\mathsf{E},w} = 1$ if and only if $f_{\mathsf{E}}(w, [\![A]\!]^{\mathsf{E}}) \subseteq [\![B]\!]^{\mathsf{E}}$

We make four assumptions about f_{E}.

> **Success**
> $f_{\mathsf{E}}(w, [\![A]\!]^{\mathsf{E}}) \subseteq [\![A]\!]^{\mathsf{E}}$

> **Non-Vacuity**
> If $\mathsf{E}(w) \cap [\![A]\!]^{\mathsf{E}} \neq \emptyset$, then $f_{\mathsf{E}}(w, [\![A]\!]^{E}) \neq \emptyset$

> **Minimality**
> If $w \in [\![A]\!]^{\mathsf{E}}$, then $f_{\mathsf{E}}(w, [\![A]\!]^{\mathsf{E}}) \subseteq \{w\}$.

> **Epistemic Accessibility Constraint**
> $f_{\mathsf{E}}(w, [\![A]\!]^{\mathsf{E}}) \subseteq \mathsf{E}(w)$

Let me briefly say what each of these constraints does for us. Success secures the validity of Identity, the principle that ⌜A > A⌝ is always true. Non-Vacuity secures a form of Conditional Non-Contradiction: if A is epistemically live, ⌜A > B⌝ and ⌜A > ¬B⌝ cannot both be true. Minimality ensures that the indicative ⌜A > B⌝ entails the material conditional ⌜¬A or B⌝. I state this principle below because it will play an important role in predicting that Transitivity, Contraposition, and Antecedent Strengthening are pseudo-valid.

> **If-to-Or**
> A > B ⊨ ¬A or B

The Epistemic Accessibility Constraint is motivated by the feeling that indicative conditionals have an epistemic flavor—they seem to talk about possibilities that are compatible with our knowledge. Following Stalnaker (1975), we can give a more concrete argument for the Epistemic Accessibility Constraint by appealing to the following well-known principle.

> **Or-to-If**
> A or B ⊨ ¬A > B

Or-to-If is a compelling principle. Consider:[19]

(39) Either the butler or the gardener did it.

(40) Therefore, if the butler didn't do it, it was the gardener.

This inference seems excellent. Nevertheless, Or-to-If is not valid. For if it were, indicatives would be logically equivalent to material conditionals. But the overwhelming consensus of contemporary philosophers and linguists is that the material conditional analysis is inadequate.

We must find another explanation, then, of why Or-to-If is so compelling. Stalnaker's explanation is that, given the Epistemic Accessibility Constraint, Or-to-If is a *reasonable inference*: roughly, if ⌜A or B⌝ can be felicitously asserted in a given context, then if ⌜A or B⌝ is known in that context, ⌜¬A > B⌝ is also known.[20] The explanation that I will give is that Or-to-If—like Transitivity, Contraposition, and Antecedent Strengthening, as we will soon see—is pseudo-valid.[21]

[19]The example is from Stalnaker (1975).

[20]Note that this is not Stalnaker's exact explanation. Stalnaker's own theory of indicatives does not make use of an epistemic accessibility relation; instead, it uses a *context set*—a set of worlds consistent with the *common commitments* of the conversational participants. Correspondingly, Stalnaker's definition of reasonable inference is stated in terms of *acceptance*, not knowledge: if ⌜A or B⌝ can be felicitously asserted in a given context, then if ⌜A or B⌝ is accepted in that context, the conditional ⌜¬A > B⌝ is also accepted. The differences between Stalnaker's own definition of reasonable inference and the one given in the main text do not matter for my purposes.

[21]As I said in footnote 2, pseudo-validity is closely related to *quasi-validity*. (See Kolodny and

Pseudo-Validity
The argument from $A_1, A_2, \ldots, A_n$ to C is pseudo-valid if and only if, whenever the premises can be felicitously asserted and ⌜Must A_1⌝, ⌜Must A_2⌝, . . . , and ⌜Must A_n⌝ are all true, C is also true.

To say that an argument is pseudo-valid is to say that if the premises can be felicitously asserted, and are entailed by the information in a given context, then the conclusion is also true in that context.

Pseudo-valid arguments tend to strike us as excellent arguments. Or-to-If is one example. But there are others, such as the following due to Bledin (2015).

(41) Either Mrs. White did it or Miss Scarlet did it.

(42) Miss Scarlet didn't do it.

(43) Mrs. White must have done it.

And here's an example that is closely related to Or-to-If. I say:

(39) Either the butler or the gardener did it.

You infer:

(44) Therefore, the butler might have done it, and if he didn't, it was the gardener.

Neither of these arguments is classically valid. In each case, the argument can have true premises and a false conclusion if the speaker doesn't know that the premises are true. Still, both arguments seem excellent.

A natural explanation appeals to pseudo-validity. Take the second argument, for example. On any plausible way of filling in the details of the example, the premise is not just true—it is also known by the speaker. If I assert (39), then ordinarily I am doing so because I know that either the butler or the gardener did it and I don't know which. (In general, it is not appropriate to assert a disjunction unless you think both disjuncts are possible.[22]) If you accept my assertion, then you will come to know that either the butler or the gardener did it, and you will also come to know that the gardener might have done it. If the inference from (39) to (44) is pseudo-valid, then it follows that (44) is true in our context. This means that when you go on to assert (44) you say something that is true. In short, pseudo-valid arguments tend to strike as as excellent arguments because, ordi-

Macfarlane (2010) and Dorr & Hawthorne (ms).) An inference from $A_1, A_2, \ldots, A_n$ to C is quasi-valid iff C is true whenever ⌜Must A_1⌝, ⌜Must A_2⌝, . . . , and ⌜Must A_n⌝ are. Dorr & Hawthorne use quasi-validity to explain Or-to-If, as well as other attractive principles. See footnote 23.

[22]This can be justified on Gricean grounds as Stalnaker (1975) suggests.

narily, if we come to know the premises *on the basis of a successful assertion of the premises*, the conclusions are true, too.

5.2 Pseudo-Validity and the Inferences

We can now show that Transitivity, Contraposition, and Antecedent Strengthening are pseudo-valid on the Stalnakerian view I have presented.[23] We rely on the pseudo-validity of Or-to-If and the classical validity of If-to-Or. So that we have everything in front of us, I repeat both principles below. (I use '$\rightsquigarrow$' for pseudo-validity and '$\vDash$' for classical validity.)

> **Or-to-If**
> A or B $\rightsquigarrow$ ¬A > B
>
> **If-to-Or**
> A > B $\vDash$ ¬A or B

Consider Transitivity. Suppose that ⌜A > B⌝ and ⌜B > C⌝ can be felicitously asserted in a given context. Suppose further that ⌜Must (A > B)⌝ and ⌜Must (B > C)⌝ are true in that context. By If-to-Or, it follows that ⌜Must (¬A or B)⌝ and that ⌜Must (¬B or C)⌝ are both true. Transitivity is valid for the material conditional. So ⌜Must (¬A or C)⌝ is true. By Or-to-If, it follows that ⌜A > C⌝ is true. The proofs for Antecedent Strengthening and Contraposition are similar.

The fact that Transitivity and Contraposition are pseudo-valid explains why we find the inferences compelling. Ordinarily, if we come to know the premises of these inferences on the basis of a successful assertion of those premises, their conclusions are also true in our context.

We can also appeal to pseudo-validity to explain the asymmetry between forward and reverse Sobel sequences. The explanation is similar to the strict theorist's explanation. Forward Sobel sequences are often felicitous because the context tends to change midway through the sequence. To see this, suppose I assert (29), repeated below.

(29) If Alice goes to the party, it will be fun.

If you accept my assertion, then all epistemically accessible worlds where Alice goes to party will be worlds where it is fun. If Bob does not go to the party in any of these worlds, the presupposition of (30), repeated below, won't be satisfied.

[23] Stalnaker (1975) shows that Transitivity and Contraposition are reasonable inferences for indicative conditionals. (He does not discuss Antecedent Strengthening.) Dorr & Hawthorne (ms) say all three inference patterns are quasi-valid for indicatives. None of these authors extend these pragmatic accounts of Contraposition, Transitivity, and Antecedent Strengthening to counterfactuals.

(30) If Alice and Bob go to the party, it won't be fun.

And so asserting (30) will tend to change the context. The set of epistemically accessible worlds will expand until it includes some where Alice and Bob go to the party. If the party is not fun in any of these worlds, (30) will come out false in our new context.

Reverse Sobel sequences, on the other hand, are typically infelicitous because the context tends not to change when the sentences are uttered in reverse order. Consider any context in which (29) can be felicitously asserted. If Antecedent Strengthening is pseudo-valid, then (29) and (30) cannot both be *known* in any such context. Suppose, then, that I assert (30) in such a context and that I say something true. If you accept my assertion, then the presupposition of (30) must be satisfied. This, in turn, means that the presupposition of (29) is satisfied. You have no reason to choose a different accessibility relation to interpret (29). And so we should expect that when I go on to assert (29), you will be puzzled—I am asking you to accept something that you cannot accept if you have accepted (30).

6 Counterfactuals and Fake Tense

To secure the pseudo-validity of Transitivity, Contraposition, and Antecedent Strengthening for indicatives, I appealed to the close connection between indicatives and epistemic modals. Specifically, I relied on the pseudo-validity of Or-to-If and the classical validity of If-to-Or. To secure the pseudo-validity of these inferences for counterfactuals, I will appeal to a close connection between counterfactuals and the modal 'had to' (and its dual 'could have'). I will use two specific principles. First, I will rely on the classical validity of a counterfactual counterpart of Or-to-If. Second, I will rely on the pseudo-validity of a counterfactual counterpart of If-to-Or. Letting $\Box_{\mathsf{H}}$ stand for *It had to be the case that*. . . , I state the principles below.

Counterfactual Or-to-If
$\Box_{\mathsf{H}}(\text{A or B}) \vDash \neg \text{A} \mathrel{\Box\!\!\rightarrow} \text{B}$

Counterfactual If-to-Or
$\text{A} \mathrel{\Box\!\!\rightarrow} \text{B} \rightsquigarrow \Box_{\mathsf{H}}(\neg \text{A or B})$

In the next two sections, I develop a new Stalnakerian semantics for counterfactuals, and I show how it secures the classical validity of Counterfactual Or-to-If (§6) and the pseudo-validity of Counterfactual If-to-Or (§7).

I will begin in §6.1 by discussing the counterfactual ingredients of counterfactuals. We will see that counterfactuals contain past tense morphology that appears not to be interpreted in the usual temporal way. I'll discuss the two main to this so-called *fake past*—the *past-as-past* approach and the *past-as-modal* approach—and present my own theory following in the past-as-modal tradition. In §6.2, I apply my semantics for the modal past to the modals 'had to' and 'could have'—which I'll call *counterfactual modals*.

6.1 A Stalnakerian Past-as-Modal Theory of Counterfactuals

Consider the following two conditionals.

(45) If Milo is not in Seattle now, he will go for Christmas.

(46) If Milo <u>were not</u> in Seattle now, he <u>would</u> go for Christmas.

The antecedent and consequent of (46) take a morphologically past tense form. The antecedent contains 'were' instead of 'is', and the consequent contains 'would' instead of 'will'. Yet (46) does not seem to be about events in the past. For this reason, Iatridou (2000) refers to the past tense morphology in counterfactuals as *fake past*. Across a wide range of languages, she and others observe, past tense markers in counterfactuals do not seem to refer to the past; instead, they seem to have a modal meaning. (46), for example, says something about a hypothetical scenario in which Milo is not in Seattle—namely, that in that scenario, he goes to Seattle for Christmas.

How, exactly, do past tense markers in counterfactuals give rise to this modal meaning? There are two main theories. The *past-as-past* theory says that the past tense markers in counterfactuals like (46) are—contrary to appearances—interpreted in the usual temporal way. True, the antecedent and consequent of (46) do not concern Milo's past locations. But past-as-past theorists say that the past tense takes wide scope over the conditional and shifts its evaluation time to the past. On a strict implementation of this theory, for example, the past in (46) takes us back to a time when it was still historically possible for Milo not to be in Seattle, and tells us that, at the time, it was necessary that either Milo would be in Seattle now or for Christmas.[24]

The second *past-as-modal* theory says that the past tense does indeed have a modal interpretation in counterfactuals like (46). According to Iatridou, the past tense has a core schematic meaning expressing some kind of *distance*. In its temporal interpretation, the past conveys *temporal distance* from a designated time.

[24] See Ippolito (2013), Khoo (2015), and Khoo (2022) for past-as-past theories.

In its modal interpretation, the past conveys *modal distance* from a designated modal parameter, such as a world or a set of worlds.[25]

The theory that I favor follows in the past-as-modal tradition. I assume that a counterfactual involves a past tense operator scoping over an indicative conditional.

(47) Past [A > B]

I assume that the meaning of an indicative conditional is given by Stalnaker's semantics. I propose that the past tense, in its modal interpretation, shifts the accessibility relation relative to which we interpret the embedded conditional. Building on suggestions in Heim (1992), Schulz (2014), and Mackay (2019), I say, roughly, that a counterfactual ⌜A $\Box\!\!\rightarrow$ B⌝ is true relative to our actual information if and only if the corresponding indicative ⌜A > B⌝ is true relative to a contextually-determined *weakening* of our information.[26]

I assume a *referential* theory of tense. On this theory, tenses are like pronouns.[27] Just as the pronoun 'she' refers to a particular, contextually-salient individual, the past tense—in its temporal interpretation—refers to a particular, contextually-salient time. Formally, tenses are indexed to free variables whose values are determined by a contextually-supplied assignment function g. ⌜Past_i A⌝ says that A is true at the time picked out by g. Following Heim (1994), I assume that, in its temporal interpretation, the past carries a presupposition of *temporal precedence*: ⌜Past_i A⌝ presupposes that $g(i)$ precedes the time of utterance.

What does a referential theory of the *modal* past look like? I will say that, just as the temporal past refers to a particular, contextually-salient time, the modal

[25]See Iatridou (2000), Schulz (2014), and Mackay (2019) for past-as-modal theories.

[26]Heim (1994) is interested in presuppositions in the antecedents of counterfactuals. She observes that when it is common ground that Mary attended the party, both (48) and (49) can be licensed.

(48) If John attended too, . . .

(49) If John had attended too, . . .

On the basis of examples like this, she hypothesizes that 'the antecedent of the counterfactual is not really added to an *empty* context, but to one which is in some sense a *revision* of the common ground *c*.' Schulz (2014) says that counterfactuals are interpreted relative to an *epistemic domain* $\mathcal{E} = \langle W, E \rangle$ where W is a set of worlds and E is a set of sets of worlds. She shows how we can reconstruct $\mathcal{E}$ from a structure that is completely parallel to a *temporal* structure $\mathcal{T}$, and then argues that we can obtain the meaning of the modal past by translating the meaning of the temporal past to the epistemic domain. I believe that Schulz's theory can also be implemented in my framework (which uses epistemic accessibility relations instead of sets of worlds). The view that I develop in the main text is also very similar in spirit and implementation to that of Mackay (2019). The main difference is that Mackay is working with a Kratzerian theory of conditionals—on which *if*-clauses function as *restrictors* on covert necessity modals—rather than a Stalnakerian theory.

[27]See Partee (1973).

past refers to a particular, contextually-salient information state—an accessibility relation. Context supplies a variable assignment g that assigns an accessibility relation to each free variable. In its modal interpretation, ⌜Past_i A⌝ says that A is true relative to the accessibility relation picked out by g—$g(i)$. Formally, where w is a world, g is a variable assignment, and E is an accessibility relation representing what's known by the conversational participants, we have:

Semantics for Modal Past
$[\![\text{Past}_i\ \text{A}]\!]^{w,g,\mathsf{E}} = 1$ if and only if $[\![\text{A}]\!]^{w,g,g(i)} = 1$

The presupposition of temporal precedence is replaced by a presupposition of *informational precedence*. Say that one accessibility relation E_1 is less informed than another accessibility relation E_2 (in symbols: $\mathsf{E}_1 < \mathsf{E}_2$) if and only if $\mathsf{E}_2(w) \subset \mathsf{E}_1(w)$ for all w. In its modal interpretation, ⌜Past_i A⌝ presupposes that the accessibility relation $g(i)$ is less informed than E.[28]

Combining the semantics for the modal past with Stalnaker's semantics for conditionals gives us the following semantics for counterfactuals.

Stalnakerian Semantics for Counterfactuals
$[\![\text{A} \mathrel{\Box\!\!\rightarrow} \text{B}]\!]^{w,g,\mathsf{E}} = [\![\text{Past}_i\ [\text{A} > \text{B}]]\!]^{w,g,\mathsf{E}} = 1$ iff $f_{g(i)}(w, [\![\text{A}]\!]^{g,g(i)}) \subseteq [\![\text{B}]\!]^{g,g(i)}$

This says that a counterfactual ⌜A $\mathrel{\Box\!\!\rightarrow}$ B⌝ is true at a world w, relative to our information state E, just in case the corresponding indicative conditional ⌜A > B⌝ is true at w, relative to a contextually-determined, less informed accessibility relation $g(i)$. I will call this accessibility relation a *counterfactual accessibility relation*.

It is well established that in evaluating a counterfactual we hold fixed certain facts about the world and not others. We can think of $g(i)$ as representing the information we're holding fixed. Take an example. We know Matt made it on time to the dinner at six. We're told he caught the bus at five. Doubting that leaving at five left him enough time to get here by six, you say:

(50) If Matt had left at five, he wouldn't have made it to the dinner on time.

To evaluate (50), we temporarily suspend some of our knowledge—our knowledge of the fact that Matt made it to the dinner on time, among other things. But we hold much of what we know fixed. In particular, we tend to hold fixed much

[28]See Mackay (2019) for a similar suggestion. Mackay models information states as sets of propositions (which, in turn, determine sets of worlds), but the main idea is the same—that the role of the modal past is to shift the information state that we use to interpret the conditional to a less informed information state.

of our knowledge about what happened before the time of the events described in the antecedent. We hold fixed when Matt started to get dressed, which buses were running at that time, and so forth. What we do and do not hold fixed is represented by the accessibility relation $g(i)$. If we're holding fixed facts about when Matt started to get dressed, then $g(i)$ takes each world w to a set of worlds consistent with what we know, in w, about when he started getting dressed. If we're holding fixed facts about the bus schedules, then $g(i)$ takes each world w to a set of worlds that is consistent with what we know, in w, about the bus schedules.

When we introduced Stalnaker's semantics for indicatives in §5, we said that there are four constraints on the selection function. For indicatives, the selection function is indexed to E—the accessibility relation representing what the conversational participants know. For counterfactuals, the selection function is indexed to $g(i)$—the accessibility relation representing what we're holding fixed for the purpose of evaluating counterfactuals. Here are the four constraints.

Success
$f_{g(i)}(w, [\![\text{A}]\!]^{g,g(i)}) \subseteq [\![\text{A}]\!]^{g,g(i)}$

Non-Vacuity
If $g(i)(w) \cap [\![\text{A}]\!]^{g,g(i)} \neq \emptyset$, then $f_{g(i)}(w, [\![\text{A}]\!]^{g,g(i)}) \neq \emptyset$

Minimality
If $w \in [\![\text{A}]\!]^{g,g(i)}$, then $f_{g(i)}(w, [\![\text{A}]\!]^{g,g(i)}) \subseteq \{w\}$.

Counterfactual Accessibility Constraint
$f_{g(i)}(w, [\![\text{A}]\!]^{g,g(i)}) \subseteq g(i)(w)$

Success secures the validity of Identity for counterfactuals: ⌜A $\Box\!\!\rightarrow$ A⌝ is always true. Non-Vacuity secures a form of Counterfactual Non-Contradiction: if there are counterfactually accessible A-worlds, ⌜A $\Box\!\!\rightarrow$ B⌝ and ⌜A $\Box\!\!\rightarrow$ ¬B⌝ cannot both be true. Minimality secures the validity of Modus Ponens.

The Counterfactual Accessibility Constraint says that the selected antecedent-world must be counterfactually accessible. Return to the dinner case. Doubting that leaving at five left him enough time to get here by six, you say:

(50) If Matt had left at five, he wouldn't have made it to the dinner on time.

I said that what we do and do not hold fixed is represented by the accessibility relation $g(i)$. If we're holding fixed facts about when Matt started to get dressed, then, as I said earlier, $g(i)$ takes each world w to a set of worlds consistent with what we know, in w, about when he started getting dressed. In that case, the

Accessibility Constraint says that the selected antecedent-world must be consistent with what we know, in w, about when he started to get dressed. If we're holding fixed facts about the bus schedules, then $g(i)$ takes each world w to a set of worlds that is consistent with what we know, in w, about the bus schedules. In that case, the Accessibility Constraint says that the selected antecedent-world must be consistent with what we know, in w, about the bus schedules.

6.2 Counterfactual Modals

We now have a theory of the modal past in place. I have used that theory to give a semantics for counterfactuals. In this section, I turn to counterfactual modals. I observe that 'had to' and 'could have' contain most past tense morphology that is similar to the past tense morphology we find in counterfactuals, and I will use my theory of the modal past to give a semantics for these counterfactual modals.

Take an example. It's raining outside. Our roommate David comes home, and leaves his wet umbrella at the door. You and I are heading out for a walk. Just as we step into the rain, I notice that you aren't wearing your raincoat.

(51) Why didn't you wear your raincoat? It had to have been raining out right now—you saw David's umbrella.

Or suppose instead that it is no longer raining by the time we get outside. Surprised that you didn't prepare for rain, I say:

(52) Why didn't you wear your raincoat? It could have been raining out right now—you saw David's umbrella.

In both (51) and (52), we find past tense morphology on both counterfactual modals. We have 'had to be raining' instead of 'has to be raining' in the case of (51) and 'could have been raining' instead of 'could be raining' in the case of (52). (51) seems to be saying, roughly, that we knew, or should have known, that it would be raining when we stepped outside; likewise, (52) seems to be saying, roughly, that rain was compatible with an impoverished knowledge state—in this case, our knowledge state before stepping outside.

The idea that (51) and (52) involve fake past has not been explored. But, as we will see in a moment, we have good reason to believe that counterfactuals are interpreted uniformly with 'had to' and 'could have'. If that's right, then if we decide on a past-as-modal approach to counterfactuals—as I have suggested—it is natural to extend this approach to counterfactual modals.[29]

[29]Dorr & Hawthorne (ms) independently observe that there is a close connection between counterfactuals and the modal 'could have' and defend a version of Counterfactual Or-to-If.

I assume that the modal claim ⌜It had to be that A⌝ involves a past tense operator scoping over the modal 'has to'.

(53) Past [has to be that A]

I will assume that 'has to' has an epistemic interpretation with the same meaning as epistemic 'must'. And I will continue to assume a standard relational semantics for epistemic modals, restated below with our new semantic parameters.

Epistemic Modals
$[\![\text{Has to be that A}]\!]^{w,g,\mathsf{E}} = [\![\text{Must A}]\!]^{w,g,\mathsf{E}} = 1$ if and only if $\mathsf{E}(w) \subseteq [\![\text{A}]\!]^{g,\mathsf{E}}$

Combining our semantics for the modal past with our relational semantics for epistemic modals gives us the following semantics for counterfactual modals.

Counterfactual Modals
$[\![\text{Had to be that A}]\!]^{w,g,\mathsf{E}} = [\![\text{Past}_i \text{ [has to be that A]}]\!]^{w,g,\mathsf{E}} = 1$ iff $g(i)(w) \subseteq [\![\text{A}]\!]^{g,g(i)}$

This says that the counterfactual modal ⌜It had to have been that A⌝ is true at a world w, relative to our information state E, just in case the epistemic modal claim ⌜It has to be that A⌝ is true at w, relative to a contextually-determined counterfactual accessibility relation $g(i)$. I assume that 'could have' is the dual of 'had to' and therefore that ⌜It could have been that A⌝ is true just in case there is some A-world that is counterfactually accessible from w.

Take the raincoat example. As we step into the rain, I say:

(51) Why didn't you wear your raincoat? It had to have been raining out right now—you saw David's umbrella.

On my view, (51) is true at a world w, and relative to our actual information E, if and only if the epistemic modal claim

(54) It has to be raining.

is true relative to a contextually-determined counterfactual accessibility relation $g(i)$. In the context most naturally evoked by (51), this counterfactual accessibility relation represents what we were in a position to know just before stepping outside. Thus, while (54) says that all worlds compatible with our present information—which includes, among other things, the fact that we're now seeing raindrops—are worlds where it's raining, (51) says that all worlds compatible with our earlier information—which does not include the fact that we see raindrops, but does include the fact that David was carrying an umbrella—are worlds where it's raining.

Now that we have a semantics for counterfactual modals in place, we can show that this theory, when combined with our Stalnakerian semantics for counterfactuals, secures the classical validity of Counterfactual Or-to-If.[30]

Counterfactual Or-to-If
$\Box_{\mathsf{H}}(\text{A or B}) \vDash \neg\text{A} \,\Box\!\!\rightarrow \text{B}$

This principle looks just as plausible as its indicative counterpart. Take an example adapted from Edgington (2008). We're hunting for a treasure. The organizer says, "I'll give you a hint. It's either in the attic or the garden." Trusting him, I go to the attic and tell my partner to search the garden. I discover the treasure. "Why did you tell me to search the garden?" she asks. I reply:

(55) The treasure had to have been either in the attic or the garden. (The organizer told me it was in one of those places.)

My partner concludes:

(56) If the treasure hadn't been in the attic, it would have been in the garden.

This inference seems excellent.

7 Sequence Semantics and Counterfactual If-to-Or

I said I would rely on two principles to secure the pseudo-validity of Transitivity, Contraposition, and Antecedent Strengthening for counterfactuals. We have now seen that the first principle—Counterfactual Or-to-If—is classically valid. In this section, I turn to the second principle, repeated below.

Counterfactual If-to-Or
$\text{A} \,\Box\!\!\rightarrow \text{B} \rightsquigarrow \Box_{\mathsf{H}}(\neg\text{A or B})$

I'll start in §7.1 by giving three arguments in favor of Counterfactual If-to-Or. In §7.2, I'll show that our past-as-modal theory of counterfactual conditionals and modals does not, as it stands, secure this principle. I will diagnose the problem and, in §7.3, offer a new *sequence semantics* for modals and conditionals, following van Fraassen (1976). In §7.4, I show that the resulting theory secures the pseudo-validity of Counterfactual If-to-Or, and as a result, predicts that Transitivity, Contraposition, and Antecedent Strengthening are pseudo-valid for counterfactuals.

[30] *Proof.* Suppose $[\![\text{Past}_{\text{i}}[\text{Has to be that A or B}]]\!]^{w,g,\mathsf{E}} = 1$. Then by the semantics for the modal past, $[\![\text{Has to be that A or B}]\!]^{w,g,g(i)} = 1$. It follows that $g(i)(w) \cap [\![\neg\text{A}]\!]^{g,g(i)} \subseteq [\![\text{B}]\!]^{g,g(i)}$. By Success and the Accessibility Constraint, $f_{g(i)}(w, [\![\neg\text{A}]\!]^{g,g(i)}) \subseteq g(i)(w) \cap [\![\neg\text{A}]\!]^{g,g(i)}$. By Stalnaker's Semantics $[\![\neg\text{A} > \text{B}]\!]^{w,g,g(i)} = 1$. By the semantics for modal past, $[\![\text{Past}_{\text{i}}\ [\neg\text{A} > \text{B}]]\!]^{w,g,\mathsf{E}} = [\![\neg\text{A} \,\Box\!\!\rightarrow \text{B}]\!]^{w,g,\mathsf{E}} = 1$.

7.1 Counterfactual If-to-Or

Counterfactual If-to-Or says that if a counterfactual $\ulcorner A \mathrel{\Box\!\!\rightarrow} B \urcorner$ is assertable in a given context and $\ulcorner$Must $A \mathrel{\Box\!\!\rightarrow} B \urcorner$ is true in the context, then $\ulcorner \Box_H(\neg A \text{ or } B) \urcorner$ is also true in the context. Remember, we are assuming that $\ulcorner$Must A$\urcorner$ is true in a given context if and only if A is known by the conversational participants in that context. With this assumption in place, Counterfactual If-to-Or says that if $\ulcorner A \mathrel{\Box\!\!\rightarrow} B \urcorner$ is assertable and known by the conversational participants in a given context, then $\ulcorner \Box_H(\neg A \text{ or } B) \urcorner$ is true in that context.

To flesh out what Counterfactual If-to-Or amounts to, we need to say what the assertability conditions are for counterfactuals. A tempting first thought is that a counterfactual is assertable only if its antecedent is known to be false. But there are well-known counterexamples to this generalization. A patient enters the emergency room displaying symptoms of what the doctor suspects is arsenic poisoning.[31] The doctor says:

(57) If the patient had taken arsenic, he would have been showing exactly the symptoms he is actually showing.

The doctor does not believe that the patient did not take arsenic—indeed, (57) is most naturally interpreted as evidence that the patient did take arsenic. In light of examples like these, we should not say that a counterfactual is assertable only if its antecedent is known to be false.

But a weaker generalization is plausible—that a counterfactual is assertable only if its antecedent is not known to be true.[32] Notice that if the doctor and his interlocutors know that the patient took arsenic, (57) is no longer acceptable. Similar things can be said about other examples.

(58) David is at the party.

(59) #If David had come to the party, he would have given a speech.

In the remainder of the paper, I will assume for simplicity that this is the only licensing condition for counterfactuals. This assumption is not essential for my main arguments. It is made purely for ease of exposition: if we say that a counterfactual is licensed if and only if its antecedent is now known, then we can treat Counterfactual If-to-Or as equivalent to the principle that whenever the counterfactual $A \mathrel{\Box\!\!\rightarrow} B$ is known, and its antecedent is not, $\ulcorner \Box_H(\neg A \text{ or } B) \urcorner$ is true.

In the remainder of this section, I provide three arguments for Counterfactual If-to-Or. The first argument is that is never acceptable to assert the *if*-claim—

[31] Anderson (1951).

[32] See von Prince (2019) for a defense of this claim. See also Starr (2014) for similar examples.

that is, the counterfactual $\ulcorner A \; \Box\!\!\rightarrow \; B\urcorner$—while denying the necessity of the *or*-claim—that is, while asserting that it could have been that A and ¬B. Take Edgington's treasure case. I say:

(56) If the treasure hadn't been in the garden, it would have been in the attic.

If you trust me and accept (56), you must also accept:

(60) The treasure couldn't have been hidden in the kitchen.

The conjunction (61) is completely unacceptable.

(61) #If the treasure hadn't been in the garden, it would have been in the attic. But it could have been in the kitchen.

Take another example. Doubting that leaving at five left Matt enough time to get to dinner by six, I say:

(50) If Matt had left at five, he wouldn't have made it to dinner by six.

Suppose you think an hour could have been enough time. You say:

(62) Maybe he could have left at five and made it by six.

Here you are presenting (62) as a challenge to (50), and if I'm persuaded to accept (62), I have to retract (50). Under no circumstances can I say:

(63) # If he had left at five, he wouldn't have made it by six. But maybe he could have left at five and made it by six.

The second argument is that Counterfactual If-to-Or follows from:

Duality
$A \Box\!\!\rightarrow B \;\rightsquigarrow\; \neg(A \Box\!\!\rightarrow \Diamond_H \neg B)$

The case for Duality is straightforward: conjunctions of the form $\ulcorner A \Box\!\!\rightarrow B$ and $A \Box\!\!\rightarrow \Diamond\neg B\urcorner$ are invariably defective.

(64) #If I had gotten an A on the exam, I would have failed the course. But if I had gotten an A on the exam, I could have passed the course.

(65) #If the treasure hadn't been in the garden, it wouldn't have been in the attic. But if it hadn't been in the garden, it could have been in the attic.

Duality and Counterfactual If-to-Or are closely related. Given the classical validity of Counterfactual Or-to-If, Duality entails Counterfactual If-to-Or.[33]

[33]Suppose that (1) $A \Box\!\!\rightarrow B$ is assertable and (2) $\ulcorner$Must$(A \Box\!\!\rightarrow B)\urcorner$ is true. Suppose, for contradiction, that (3) $\Diamond_H(A\wedge\neg B)$ is true. By (3) it follows that $\Diamond_H\neg B$. Assuming the counterfactual accessibility relation is an equivalence relation, it follows that $\Box_H\Diamond_H\neg B$ is true, and so $\Box_H(\neg A \vee \Diamond_H\neg B)$

The third argument for Counterfactual If-to-Or, which is closely related to the second argument, concerns a specific form of context-sensitivity of counterfactuals on which the standards for what can and cannot be held fixed can be more or less demanding. Take a case from Lewis (2016) and Gillies (2007). Sophie is considering going to a parade to see Pedro Martinez—her favorite baseball player—but decides not to. You and I go. I say:

(66) If Sophie had gone to the parade, she would have seen Pedro.

When I assert (66) I'm holding fixed all sorts of facts about the parade—that it isn't cancelled because of rain, that Pedro does not get the flu, that there are no tall people obstructing what would have been Sophie's view. Now suppose you mention the possibility of Sophie getting stuck behind a tall person. You say:

(67) She *could have* gone to the parade and gotten stuck behind a tall person (and so not seen Pedro).

Suppose I'm persuaded to accept (67). Then my standards for what facts I can hold fixed—and so for what counterfactuals I can assert—will become more demanding. I will no longer willing to hold fixed that there are no tall people obstructing what would have been Sophie's view, and so I will no longer be willing to assert (66).

How do we account for these changes in the assertability of a counterfactual? One natural answer is offered by Lewis (2016). Lewis says, roughly, that a counterfactual is true in a context just in case all of the closest antecedent-worlds that are *relevant* in that context are consequent-worlds. In the parade case, if we're ignoring remote possibilities in which Sophie is stuck behind a tall person for the duration of the parade, those possibilities are not relevant, and so (66) is true. If we start to take this possibility seriously, it becomes relevant, and (66) will be false in our new context.

But if Conditional Excluded Middle is true—and there is excellent evidence that it is, including the discussion of the probabilities of conditionals in §3—Lewis's account cannot be right. Lewis says that (66) is false because there are counterfactually accessible worlds where Sophie goes to the parade and does not see Pedro—that is, because (67) is true. If that's right, then (68) is also false.

(68) If Sophie had gone to the parade, she would not have seen Pedro dance.

After all, Sophie could have gone to the parade and seen Pedro dance. But Conditional Excluded Middle says that at least one of (66) and (68) is true.

is too. By Counterfactual Or-to-If, it follows that $A \mathrel{\Box\!\!\rightarrow} \Diamond_H \neg B$. But by Duality, (1) and (2) entail $\neg(A \mathrel{\Box\!\!\rightarrow} \Diamond_H \neg B)$. Contradiction.

Those, like me, who accept Conditional Excluded Middle cannot accept Lewis's account, then. So how should we account for the shiftiness of counterfactuals in cases like the parade case? I suggest that we appeal to Counterfactual If-to-Or. To see this, suppose you assert (67) and that I accept your assertion. Then the context changes. Our standards for what can be held fixed are now more demanding—we are no longer willing to hold fixed that there are no tall people obstructing what would have been Sophie's view. In other words, more worlds have become counterfactually accessible—specifically, some where Sophie worlds goes to the parade and gets stuck behind a tall person for the duration of the parade. (Remember, we are assuming that 'could have' is interpreted uniformly with counterfactuals.) Given Conditional Excluded Middle, the counterfactual (66), repeated below, may still be true in our context.

(66) If Sophie had gone to the parade, she would have seen Pedro.

But Counterfactual If-to-Or predicts that the conversational participants no longer count as *knowing* that (66) is true. Why? Counterfactual If-to-Or says that insofar as there are counterfactually possible worlds where Sophie goes to the parade and does not see Pedro—that is, insofar as (67) is true in our context—there are *epistemically* possible worlds in which the counterfactual (66) is false.[34]

I have now given three arguments for Counterfactual If-to-Or: first, that conjunctions of the form $\ulcorner A \mathrel{\Box\!\!\rightarrow} B$ and $\Diamond_H(A \text{ and } \neg B)\urcorner$ are invariably defective; second, that Counterfactual If-to-Or follows from Duality; and third, that it helps explain the shiftiness of counterfactuals. Taken together, these three arguments make a strong case for Counterfactual If-to-Or. But in the next section I'll show that our theory does not yet secure this principle.

7.2 The Need for Fine-Grained Contents

Counterfactual If-to-Or says that which counterfactuals we, the conversational participants, count as knowing in a given context depends, in part, on which 'could have' claims are true in our context. If, for all we know, Matt did not flip the coin, and the coin could have landed heads, then it follows that for all we know it would have landed heads if it had been flipped. To validate this principle, there must be constraints on the relationship between epistemic possibility and counterfactual possibility: if we don't know A, and there are counterfactually possible A-worlds where B is true, then there must be epistemically possible

[34]Note that I am assuming here that the conversational participants do not know that Sophie goes to the parade in the context described. See Boylan (forthcoming) for discussion of similar issues.

worlds where the counterfactual ⌜A $\Box\!\!\rightarrow$ B⌝ is true.

But nothing we have said so far guarantees that this will always be so. To see this, consider a simple model. There are just three worlds: $\{w_1, w_2, w_3\}$. In w_1, Matt flips a fair coin, and it lands heads. In w_2, the coin lands tails. In w_3, he does not flip the coin. Suppose I know that he does not flip the coin: $\mathsf{E}(w_3) = \{w_3\}$. And suppose that all three worlds are counterfactually possible: $g(i)(w_3) = \{w_1, w_2, w_3\}$. Then (69)-(71) are all true in w_3.

(69) Matt might not have flipped the coin.

(70) The coin could have landed tails.

(71) The coin could have landed heads.

By Conditional Excluded Middle, one of (72) and (73) is true in w_3.

(72) If Matt had flipped the coin, it would have landed heads.

(73) If Matt had flipped the coin, it would have landed tails.

Regardless of which of (72) or (73) is true, we have a counterexample to Counterfactual If-to-Or. If (72) is true, then I know (72)—w_3 is the only epistemically accessible world. But it's also true, in my context, that the coin could have landed tails—(70) is true. Likewise, if (73) is true, then I know (73) even though it's also true, in my context, that the coin could have landed heads.

To rule out this model, we need a certain *plenitude* assumption—an assumption to the effect that if we're at an epistemically possible ¬A-world, then every counterfactually possible A-world w is such that it is epistemically possible that if it had been that A, it would have been that w. In the simple three-world model, this plenitude assumption fails—there are two few epistemic possibilities, given which worlds are counterfactually possible. In w_3, there is a counterfactually possible world where the coin is flipped and lands heads (w_1); there is also a counterfactually possible world where the coin is flipped and lands tails (w_2). But there is only *one* epistemically possible world—either Flip $\Box\!\!\rightarrow$ Heads is epistemically possible, or Flip $\Box\!\!\rightarrow$ Tails is epistemically possible, but not both.

The problem of having too few epistemic possibilities is a familiar one in the literature on the probabilities where it is known as the *wallflower problem*. Let me introduce a simplified version of the wallflower problem found in Bacon (2015).[35] Suppose I roll a six-sided die, but I have not seen how it landed. You might have thought we can model my epistemic state with exactly six equiprobable worlds, one for each outcome of the die roll. But this won't work. To see why

[35] See Hájek (1989). My presentation follows Bacon (2015).

not, consider the indicative conditional:

(74) If it did not land on six, it landed on one.

I assign a probability of 1/5 to (74). But in our simple six-world model, there is no proposition whose probability is equal to 1/5. With just six equiprobable worlds, the unconditional probability of any proposition—and so, in particular, any conditional proposition—is a multiple of 1/6.

As Bacon (2015) convincingly explains, if we accept Conditional Excluded Middle, we should say that there's a flaw in our simple model. To see this, suppose the die lands on six. By Conditional Excluded Middle, one of the following conditionals must be true.

(75) If it landed on four or five, it landed on four.

(76) If it landed on four or five, it landed on five.

But clearly I am in no position to know *which* of (75) or (76) is true. So, Bacon says, there must be at least two epistemic possibilities compatible with the die landing on six—one where (75) is true, one where (76) is true. It is easy to see that we can generate many more epistemic possibilities by considering other antecedents that are false when the die lands on six. In short, if we accept Conditional Excluded Middle, we must countenance many more epistemic possibilities than our original six.

I say that there is a similar flaw in our model of the coin flip. I said that there were three counterfactually accessible worlds: w_1 (where the coin is tossed and lands heads), w_2 (where the coin is tossed and lands tails) and w_3 (where the coin is not tossed). And I said that, in w_3, w_3 is the only epistemic possibility—I know that the coin is not tossed in w_3. But again if we accept Conditional Excluded Middle, this can't be right. As we have seen, Conditional Excluded Middle entails that one of (72) or (73) is true in w_3.

(72) If Matt had flipped the coin, it would have landed heads.

(73) If Matt had flipped the coin, it would have landed tails.

But clearly I am in position to know which of (72) or (73) is true. So, w_3 needs to be split into two epistemic possibilities—one where (72) is true, and one where (73) is true.

These days it's common for philosophers to model these more fine-grained possibilities with *sequences* of 'factual' worlds, following van Fraassen (1976).[36] To see how these models work, return to the case of the die. Consider the possi-

[36] See, among others, Bacon (2014), and Goldstein & Santorio (2021).

bility that the die lands on six. We have seen that there are many ways of settling the indicative conditional facts that are compatible with the die landing on six—it could be that if the die didn't land on six, it landed on three, or it could be that if the die didn't land on six, it landed on two, and so forth. Each of these epistemic possibilities is modeled as a sequence of worlds. For example, consider the sequence:

$$\langle w_6, w_2, w_3, w_4, w_5, w_1 \rangle$$

This sequence represents one way of the settling all of the facts—both the non-conditional facts and the indicative conditional facts. The first world tells us how the non-conditional facts have been settled—in this case, it tells us that the die landed on six. The other worlds in the sequence tell us how the indicative conditional facts have been settled. For example, the second world tells us what is true if we are not in the first world. This sequence tells us that if the die didn't land on six, it landed on two. The third world tells us what is true if we are not in the first or the second world. This sequence tells us that if it didn't land on six or two, it landed on three. And so on.

In the next section, I suggest a generalization of van Fraassen's sequence semantics—one that provides a simple, uniform semantics for indicatives and counterfactuals. Roughly, I will say that an indicative conditional is true at a sequence just in case the first *epistemically possible* antecedent-world in the sequence is a consequent-world, and that a counterfactual is true at a sequence just in case the first *counterfactually possible* antecedent-world is a consequent-world. (Note: I will only be considering *simple conditionals* in the next section—conditionals whose antecedents and consequents do not themselves contain conditionals. There is a way to generalize the models to all conditionals, but the details would take us too far afield.)

7.3 A Uniform Sequence Semantics for Conditionals

Begin with a finite set of 'factual' worlds W. Let $\mathbf{S}_W$ be the set of all permutations of W. Thus, where the elements of W represent all possible ways of settling the non-conditional facts, the elements of $\mathbf{S}_W$ represent all possible ways of settling all of the facts—the non-conditional facts and the conditional facts alike. Where s is a sequence in $\mathbf{S}_W$, we will write w_s for the first world in s. The semantics for non-conditional sentences is simple. A non-conditional sentence A is true at a sequence s just in case A is true at w_s.

We interpret epistemic modals and indicatives relative to an epistemic accessibility relation. Before we used *factual accessibility relations*—accessibility relations over W—to interpret conditionals. Now we will need to use *lifted accessibility relations*—accessibility relations over $\mathbf{S}_W$. Where E is any factual epistemic accessibility relation over W, E determines a lifted accessibility relation $\mathbf{E}$ over $\mathbf{S}_W$ as follows.

Definition of Lifted Accessibility
$\mathbf{E}(s) = \{s' \in \mathbf{S}_W : w_{s'} \in \mathsf{E}(w_s)\}$

A sequence s_1 is $\mathbf{E}$-accessible from another sequence s_2 if and only if the first world in s_1 is E-accessible from the first world in s_2.

Now that we have lifted accessibility relations, we need to redefine our selection functions. To do so, I need to introduce one piece of terminology. For any set of sequences $\mathbf{P}$, we define the *flattening* of $\mathbf{P}$ as follows.

Flattening
$\downarrow \mathbf{P} = \{w \in W : w = w_s \text{ for some } s \in \mathbf{P}\}$

The flattening of $\mathbf{P}$ is the set of first worlds of the sequences in $\mathbf{P}$. Where $\mathbf{E}$ is a lifted accessibility relation over $\mathbf{S}_W$, we define the selection function as follows.

$$f_{\mathbf{E}}(s, \mathbf{P}) = \begin{cases} \text{the singleton of the 1st} \downarrow [\mathbf{E}(s) \cap \mathbf{P}]\text{-world in s} & \text{if } \downarrow [\mathbf{E}(s) \cap \mathbf{P}] \neq \emptyset \\ \emptyset & \text{otherwise} \end{cases}$$

The selection function $f_{\mathbf{E}}$ takes a sequence s and a set of sequences $\mathbf{P}$ to a set containing at most one world: the first E-accessible world in s where $\downarrow \mathbf{P}$ is true if there are any E-accessible worlds where $\downarrow \mathbf{P}$ is true and the empty set otherwise.

We are now in a position to state the semantics for epistemic modals and indicative conditionals. Let s be any sequence. Let $\mathbf{E}$ be a lifted epistemic accessibility relation. Let $\boldsymbol{g}$ be a *lifted variable assignment*. (I will say what this is in a moment when I introduce the semantics for the modal past.) Then we have:

Epistemic Modals
$[\![\text{Must A}]\!]^{s,\boldsymbol{g},\mathbf{E}} = [\![\text{Has to be that A}]\!]^{s,\boldsymbol{g},\mathbf{E}} = 1$ if and only if $\mathbf{E}(s) \subseteq [\![\text{A}]\!]^{\boldsymbol{g},\mathbf{E}}$

Stalnaker's Semantics
$[\![\text{A} > \text{B}]\!]^{s,\boldsymbol{g},\mathbf{E}} = 1$ if and only if $f_{\mathbf{E}}(s, [\![\text{A}]\!]^{\boldsymbol{g},\mathbf{E}}) \subseteq \downarrow [\![\text{B}]\!]^{\boldsymbol{g},\mathbf{E}}$

⌜Must A⌝ or ⌜Has to be that A⌝ is true at a sequence s just in case A is true at all sequences $\mathbf{E}$-accessible from s. ⌜A > B⌝ is true at a sequence s if and only if either there are no A-worlds that are E-accessible from w_s or the first A-world that is E-accessible from w_s is a B-world.

Let us check that all of the principles governing the logic of indicative conditionals discussed in §5 continue to hold.[37] First, observe that Conditional Excluded Middle is valid. Consider any sequence s. If there there are A-worlds epistemically accessible from w_s, then either the first accessible A-world is a world where B is true, or it is a world where ¬B is true. In the first case, ⌜A > B⌝ is true; in the second, ⌜A > ¬B⌝ is true. If there are no A-worlds accessible from w_s, both conditionals are true.

Second, Identity is valid: ⌜A > A⌝ is always true. This is because the selection function satisfies Success.

Third, Modus Ponens is valid. This is because the selection function satisfies Minimality: if $w_s \in \downarrow[\![A]\!]^{\boldsymbol{g},\mathsf{E}}$, then $f_{\mathsf{E}}(s, [\![A]\!]^{\boldsymbol{g},\mathsf{E}}) = \{w_s\}$.

Fourth, we secure a form of Conditional Non-Contradiction: if A is epistemically live, then ⌜A > B⌝ and ⌜A >¬B⌝ are not both true. This is because the selection function satisfies Non-Vacuity: if $\downarrow[\![A]\!]^{\boldsymbol{g},\mathsf{E}} \cap \mathsf{E}(w_s) \neq \emptyset$, then $f_{\mathsf{E}}(s, [\![A]\!]^{\boldsymbol{g},\mathsf{E}}) \neq \emptyset$.

Fifth, we secure the pseudo-validity of Or-to-If. If the disjunction ⌜A or B⌝ is epistemically necessary, then the indicative conditional ⌜¬A > B⌝ is true. This is because the selection function satisfies a version of the Epistemic Accessibility Constraint: $f_{\mathsf{E}}(s, [\![A]\!]^{\boldsymbol{g},\mathsf{E}}) \subseteq \mathsf{E}(w_s)$.

I turn now to the modal past. Let g be any 'factual' variable assignment that assigns factual accessibility relations to free variables. g determines a *lifted variable assignment* $\boldsymbol{g}$ that assigns to each free variable an accessibility relation over $\mathbf{S}_{\mathrm{W}}$. More precisely, the lifted variable assignment is defined as follows.

Lifted Variable Assignment
$\boldsymbol{g}(i)(s) = \{s' \in \mathbf{S}_{\mathrm{W}} : w_{s'} \in g(i)(w_s)\}$

We can now state the semantics for the modal past. Where s is a sequence, $\boldsymbol{g}$ is a lifted variable assignment, and E is a lifted accessibility relation, we have:

Modal Past
$[\![\mathrm{Past_i}\ A]\!]^{s,\boldsymbol{g},\mathsf{E}} = 1$ if and only if $[\![A]\!]^{s,\boldsymbol{g},\boldsymbol{g}(i)} = 1$

We continue to assume that the modal past carries a presupposition of informational precedence: ⌜Past$_i$ A⌝ presupposes that $\boldsymbol{g}(i) < \mathsf{E}$. Combining the semantics for the modal past with our semantics for epistemic modals and indicatives, respectively, gives us:

[37]Note that I am only considering *simple conditionals*—conditionals whose antecedents and consequents do not themselves contain conditionals.

Semantics for Counterfactual Modals
$[\![\text{Past}_\text{i} \text{ [Has to be that A]}]\!]^{s,\boldsymbol{g},\mathsf{E}} = 1$ if and only if $\boldsymbol{g}(i)(s) \subseteq [\![\text{A}]\!]^{\boldsymbol{g},\boldsymbol{g}(i)}$

Stalnakerian Semantics for Counterfactuals
$[\![\text{Past}_\text{i} \text{ [A > B]}]\!]^{s,\boldsymbol{g},\mathsf{E}} = 1$ if and only if $f_{\boldsymbol{g}(i)}(s, [\![\text{A}]\!]^{\boldsymbol{g},\boldsymbol{g}(i)}) \subseteq \downarrow [\![\text{B}]\!]^{\boldsymbol{g},\boldsymbol{g}(i)}$

This says that ⌜It had to be that A⌝ is true at a sequence just in case all $\boldsymbol{g}(i)$-accessible sequences are A-sequences, and that a counterfactual with ⌜A □→ B⌝ is true at a sequence just in case the first $g(i)$-accessible A-world in that sequence is a B-world. Since counterfactuals are interpreted uniformly with 'had to' and 'could have', we continue to validate Counterfactual Or-to-If. I leave the proof to a footnote.[38]

The final order of business is to check that we predict that Counterfactual If-to-Or is pseudo-valid. Earlier I said that, in order to secure the pseudo-validity of Counterfactual If-to-Or, we would need a certain plenitude assumption—an assumption that, if we're at an epistemically possible ¬A-world, then every counterfactually possible A-world w is such that it is epistemically possible that if it had been that A, it would have been that w. In the simple three-world model of the coin flip, this plenitude assumption failed—there were two few epistemic possibilities, given which worlds were counterfactually possible.

By contrast, our new sequence models respect the plenitude assumption. If we're at an A-world, then if we're searching for an A-world, we can't go beyond the first world—the truth value of a counterfactual with antecedent A is settled by the first world. But if we're at a ¬A-world, then all counterfactually possible A-worlds have some chance of being selected; that is to say, if w is a counterfactually possible A-world, then among the epistemically possible ¬A-sequences, there is at least one whose first A-world is w. This means that there are guaranteed to be epistemically possible ¬A-sequences where the conditional A □→ w is true.

This is exactly what we need to secure Counterfactual If-to-Or. I will leave the proof to a footnote, but I will give an informal explanation of how this works in the coin case.[39] We know that the coin is fair, so (70) and (71), repeated below, are both true in our context.

[38] *Proof.* Suppose $[\![\text{Past}_\text{i}\text{[Has to be that A or B]}]\!]^{s,\boldsymbol{g},\mathsf{E}} = 1$. Then by the semantics for the modal past, $[\![\text{Has to be that A or B}]\!]^{s,\boldsymbol{g},\boldsymbol{g}(i)} = 1$. It follows that $\boldsymbol{g}(i)(s) \cap [\![\neg\text{A}]\!]^{\boldsymbol{g},\boldsymbol{g}(i)} \subseteq [\![\text{B}]\!]^{\boldsymbol{g},\boldsymbol{g}(i)}$ and so also that $\downarrow [\boldsymbol{g}(i)(s) \cap [\![\neg\text{A}]\!]^{\boldsymbol{g},\boldsymbol{g}(i)}] \subseteq \downarrow [\![\text{B}]\!]^{\boldsymbol{g},\boldsymbol{g}(i)}$. By the definition of the selection function, $f_{\boldsymbol{g}(i)}(s, [\![\neg\text{A}]\!]^{\boldsymbol{g},\boldsymbol{g}(i)}) \subseteq \downarrow[\boldsymbol{g}(i)(s) \cap [\![\neg\text{A}]\!]^{\boldsymbol{g},\boldsymbol{g}(i)}] \subseteq \downarrow[\![\text{B}]\!]^{\boldsymbol{g},\boldsymbol{g}(i)}$. By our Stalnakerian semantics for counterfactuals, it follows that $[\![\text{Past}_\text{i} \text{ [}\neg\text{A > B]}]\!]^{w,\boldsymbol{g},\mathsf{E}} = [\![\neg\text{A □→ B}]\!]^{w,\boldsymbol{g},\mathsf{E}} = 1$.

[39] *Proof.* Suppose (1) $[\![\text{Might}\neg\text{A}]\!]^{s,\boldsymbol{g},\mathsf{E}} = 1$ and (2) $[\![\text{Must (Past}_\text{i} \text{ (A>B))}]\!]^{s,\boldsymbol{g},\mathsf{E}} = 1$. Suppose, for contradiction, that (3) $[\![\text{Past}_\text{i} \text{ (Could (A and }\neg\text{B))}]\!]^{s,\boldsymbol{g},\mathsf{E}} = 1$. By (3) and the semantics for modal past, there's an $s_1 \in \boldsymbol{g}(i)(s)$ such that $[\![\text{A and }\neg\text{B}]\!]^{s_1,\boldsymbol{g},\boldsymbol{g}(i)} = 1$. It follows that $w_{s_1} \in\downarrow [\boldsymbol{g}(i)(s) \cap [\![\text{A}]\!]^{\boldsymbol{g},\boldsymbol{g}(i)}]$ and $w_{s_1} \in\downarrow [\![\neg\text{B}]\!]^{\boldsymbol{g},\boldsymbol{g}(i)}$. Note that $\mathsf{E}(s)$ is the set of all permutations of W beginning with a world in $\mathsf{E}(w_s)$. Since $[\![\text{Might }\neg\text{A}]\!]^{s,\boldsymbol{g},\mathsf{E}} = 1$, there's an $s_2 \in \mathsf{E}(s)$ such that (a) $w_{s_2} \notin\downarrow$

(70) The coin could have landed tails.

(71) The coin could have landed heads.

There are counterfactually possible worlds the coin lands heads, and there are counterfactually possible worlds where the coin lands tails. The plenitude assumption guarantees that, among the the epistemically possible sequences where the coin is not tossed, there are some whose first counterfactually possible toss-world is a heads-world and some whose first counterfactually possible toss-world is a tails-world. And so we predict, in accordance with Counterfactual If-to-Or, that I do not know either of the following counterfactuals.

(72) If Matt had flipped the coin, it would have landed heads.

(73) If Matt had flipped the coin, it would have landed tails.

7.4 Pseudo-Validity and the Inferences

We now have in place a Stalnakerian theory of counterfactuals on which Counterfactual Or-to-If is classically valid and Counterfactual If-to-Or is pseudo-valid. So that we have everything in front of us, I repeat both principles below.

Counterfactual Or-to-If
$\Box_{\mathsf{H}}$(A or B) $\vDash$ ¬A $\Box\!\!\rightarrow$ B

Counterfactual If-to-Or
A $\Box\!\!\rightarrow$ B $\rightsquigarrow$ $\Box_{\mathsf{H}}$(¬A or B)

These two principles suffice to secure the pseudo-validity of our three inferences. Take Transitivity. Suppose ⌜A $\Box\!\!\rightarrow$ B⌝ and ⌜B $\Box\!\!\rightarrow$ C⌝ can be felicitously asserted in a given context and that ⌜Must (A $\Box\!\!\rightarrow$ B)⌝ and ⌜Must (B $\Box\!\!\rightarrow$ C)⌝ are both true. By the psuedo-validity of Counterfactual If-to-Or, it follows that ⌜$\Box_{\mathsf{H}}$(¬A or B)⌝ and ⌜$\Box_{\mathsf{H}}$(¬B or C)⌝ are both true. Transitivity is valid for the material conditional. So ⌜$\Box_{\mathsf{H}}$(¬A or C)⌝ is true. By the classical validity of Counterfactual Or-to-If, it follows that the ⌜A $\Box\!\!\rightarrow$ C⌝ is true. The proofs for Antecedent Strengthening and Contraposition are similar.

The fact that Transitivity and Contraposition are pseudo-valid explains why we find the inferences compelling. Ordinarily, if we come to know the premises

$[\![A]\!]^{\boldsymbol{g},\boldsymbol{g}(i)}$ and (b) w_{s_1} is the second world in s_2. Since $w_{s_1} \in\downarrow [\boldsymbol{g}(i)(s) \cap [\![A]\!]^{\boldsymbol{g},\boldsymbol{g}(i)}]$, and $\boldsymbol{g}(i)$ is an equivalence relation, $w_{s_1} \in\downarrow [\boldsymbol{g}(i)(s_2) \cap [\![A]\!]^{\boldsymbol{g},\boldsymbol{g}(i)}]$. And since w_{s_1} is the second world in s_2 and $w_{s_2} \notin [\![A]\!]^{\boldsymbol{g},\boldsymbol{g}(i)}$, w_{s_1} is the first $\downarrow [\boldsymbol{g}(i)(s_2) \cap [\![A]\!]^{\boldsymbol{g},\boldsymbol{g}(i)}]$-world in s_2. So $f_{\boldsymbol{g}(i)}(s_2, [\![A]\!]^{\boldsymbol{g},\boldsymbol{g}(i)}) = \{w_{s_2}\}$. Since $w_{s_2} \in [\![\neg B]\!]^{\boldsymbol{g},\boldsymbol{g}(i)}$, $f_{g(i)}(s_2, [\![A]\!]^{\boldsymbol{g},\boldsymbol{g}(i)}) \not\subseteq [\![B]\!]^{\boldsymbol{g},\boldsymbol{g}(i)}$. By Stalnaker's Semantics $[\![A > B]\!]^{s_2,\boldsymbol{g},\boldsymbol{g}(i)} = 0$. By the semantics for the modal past, $[\![\text{Past}_i\ (A > B)]\!]^{s_2,\boldsymbol{g},\mathsf{E}} = 0$. By the semantics for epistemic modals, $[\![\text{Must}\ (\text{Past}_i\ (A>B))]\!]^{s,\boldsymbol{g},\mathsf{E}} = 0$. But that contradicts (2).

of these inferences on the basis of a successful assertion of those premises, their conclusions must also be true in our context.

And we can appeal to pseudo-validity to explain the asymmetry between forward and reverse Sobel sequences for counterfactuals. Just as we saw with indicative conditionals, the explanation in terns of pseudo-validity is similar to the strict theorist's explanation. Forward Sobel sequences are often felicitous because the context tends to change midway through the sequence. Reverse Sobel sequences are typically infelicitous because the context tends not to change when the sentences are uttered in reverse order.

8 Conclusion

Our ordinary, non-extreme credences in conditionals differ drastically from those predicted by the strict theory. In contrast, Stalnaker's variably strict theory fares much better. Therefore, we should reject the strict theory in favor of a Stalnakerian variably strict theory. But I have argued that certain inference patterns that are valid on the strict theory and invalid on Stalnaker's theory—such as Transitivity, Contraposition, and Antecedent Strengthening—share certain attractive features with classically valid arguments. In this paper, I developed a Stalnakerian variably strict theory on which these characteristically strict inference patterns are pseudo-valid for both indicative conditionals and counterfactuals.

9 References

Adams, Ernest (1975). *The Logic of Conditionals*. Dordrecht, Holland: D. Reidel.

Bacon, Andrew (2015). "Stalnaker's thesis in context." *Review of Symbolic Logic*, 8 (1):131-163.

Bledin, Justin (2014). "Logic Informed." *Mind*, 123 (490): 277-316.

Boylan, David (forthcoming). 'Counterfactual Skepticism Is (Just) Skepticism.' *Philosophy and Phenomenological Research*.

Dorr, Cian (2016). "Against Counterfactual Miracles." *Philosophical Review*, 125 (2): 241-286.

Dorr, Cian & John Hawthorne (2018). *If...: A theory of conditionals*. Manuscript.

Edgington, Dorothy (1995). "On conditionals." *Mind*, 104 (414):235–329.

Gillies, Thony (2007). "Counterfactual Scorekeeping." *Linguistics and Philosophy*, 30 (3): 329–360.

Gillies, Thony (2004). "Epistemic Conditionals and Conditional Epistemics." *Noûs*, 30 (3): 329–360.

Goodman, Nelson (1947). "The problem of counterfactual conditionals." *Journal of Philosophy*, 44 (5):113-128.

Hájek, Alan (1989). "Probabilities of conditionals—revisited." *Journal of Philosophical Logic*, 18 (4):423 - 428.

Heim, Irene (1992). "Presupposition projection and the semantics of attitude verbs." *Journal of Semantics*, 9(3): 183–221

Heim, Irene (1994). "Comments on Abusch's Theory of Tense." In Hans Kamp (ed.), *Ellipses, Tense, and Questions*: 143-170

Herzberger, Hans (1973). "Dimensions of truth." *Journal of Philosophical Logic*, 2 (4): 535 - 556.

Holguín, Ben (2021). "Indicative Conditionals Without Iterative Epistemology." *Noûs*, 55 (3): 560–80.

Iatridou, Sabine (2000). 'The Grammatical Ingredients of Counterfactuality.' *Linguistic Inquiry* 31(2): 231–270.

Ippolito, Michela (2013). 'Subjunctive Conditionals: A Linguistic Analysis.' *Linguistic Inquiry Monograph* (Series 65), Cambridge: MIT Press.

Karttunen, Lauri & Stanley Peters (1979). "Conventional implicature." *Syntax and Semantics*, 11: 1–56.

Kaufmann, Stefan (2004). 'Conditioning against the grain: Abduction and indicative conditionals.' *Journal of Philosophical Logic* 33, pp. 583-606.

Kaufmann, Stefan (2009). 'Conditionals right and left: Probabilities for the whole family.' *Journal of Philosophical Logic* 38, pp. 1-53.

Khoo, Justin (2015). 'On Indicative and Subjunctive Conditionals.' *Philosopher's Imprint* 15, pp. 1-40.

Kolodny, Niko & MacFarlane, John (2010). 'Ifs and Oughts.' *Journal of Philosophy* 107 (3):115-143.

Khoo, Justin (2015). 'On Indicative and Subjunctive Conditionals.' *Philosopher's Imprint* 15, pp. 1-40.

Khoo, Justin (2022). 'The meaning of *If*.' Oxford University Press.

Lewis, David (1973). *Counterfactuals*. Basil Blackwell Ltd., Malden, MA.

Lowe, E. J. (1995). 'The Truth about Counterfactuals.' *Philosophical Quarterly* 45 (178):41-59.

Mackay, John (2019). "Modal Interpretation of Tense in Subjunctive Conditionals." *Semantics and Pragmatics*, 12 (2)

Mandelkern, Matthew (2018). "Talking about worlds." *Philosophical Perspectives*, 32 (1): 298-325.

Mandelkern, Matthew (2019). "Bounded Modality." *Philosophical Review*, 128 (1): 1-61.

Mandelkern, Matthew (2022). "Bounds." Manuscript

McGee, Vann (1989). 'Conditional probabilities and compounds of conditionals.' *Philosophical Review* 98, 485-541.

Partee, Barbara (1973). "Some structural analogies between tenses and pronouns in English." *Journal of Philosophy*, 70 (18): 601-609.

Santorio, Paolo (2022). "Path Semantics for Indicative Conditionals." *Mind*, 131

(521): 59-98.

Schulz, Katrin (2014). ‘Fake tense in conditional sentences: A modal approach.’ *Natural Language Semantics*

Schulz, Moritz (2017). *Counterfactuals and Probability*. Oxford: Oxford University Press.

Stalnaker, Robert (1968). ‘A Theory of Conditionals.’ In Nicholas Rescher (ed.) *Studies in Logical Theory*(American Philosophical Quarterly Monographs 2), Oxford: Blackwell. pp. 98-112.

Stalnaker, Robert (1975). ‘Indicative conditionals.’ *Philosophia* 5, pp. 269-286.

Starr, William B (2014). ‘A Uniform Theory of Conditionals.’ *Journal of Philosophical Logic* 43 (6): 1019-1064.

von Fintel, Kai (1998). ‘The presupposition of subjunctive conditionals.’ In Orin Percus & Uli Sauerland (eds.), *The interpretive tract*, Vol. 25, pp. 25- 44. Cambridge, Massachusetts: MIT Working Papers in Linguistics.

von Fintel, Kai (2001). “Counterfactuals in a Dynamic Context.” In Ken Hale: A Life in Language, Michael Kenstowicz, editor. MIT Press, Cambridge.

van Fraassen, Bas C. (1976). ‘Probabilities of conditionals.’ In *Foundations of probability theory, statistical inference, and statistical theories of science*, Springer, pp. 261 - 308.

von Prince, Kilu (2019). ‘Counterfactuality and past.’ *Linguistics and Philosophy* 42 (6):577-615.

Veltman, Frank (1985). “Logics for Conditionals.” Doctoral Dissertation, University of Amsterdam.

Warmbrod, Ken (1981). “Counterfactuals and Substitution of Equivalent Antecedents.” *Journal of Philosophical Logic*, 10: 267–289.